



SAVE-U

SENSORS AND SYSTEM ARCHITECTURE FOR VULNERABLE ROAD USERS PROTECTION



Information Society Technologies: Systems and Services for the Citizen

Project IST – 2001 – 34040

Project funded by the European Community under
the "Information Society Technology" Programme (1998-2002)

Deliverable 1-B: Vulnerable Road User Sensor Data Collection

Authors D. M. Gavrilă
(DaimlerChrysler)

Co-Authors: L. Letellier (CEA),
and M.-M. Meinecke (VW)

Dissemination Level: PU
Version: 4.0
Contract Date: 1/3/2002
Document Date: 20/01/2003

Partners: Faurecia Industries (FR), Siemens VDO Automotive AG (DE),
MIRA Ltd (GB), Commissariat à l'Énergie Atomique (FR),
DaimlerChrysler (DE), Volkswagen AG (DE)

Revision chart and history log

Version	Date	Reason
1.0	26/07/2002	First draft to SAVE-U partners for review.
2.0	19/08/2002	Second draft to SAVE-U partners for review.
2.1	28/08/2002	Complete draft to SAVE-U partners for comments
2.2	10/10//2002	Version presented at EU annual project review
3.0	15/12/2002	Complete revision according to EU review to SAVE-U partners for review
4.0	20/01/2003	Final version

Credits

This document was prepared by partners of the SAVE-U consortium. Please contact one of the following project partners for further information regarding this document or the project:

DaimlerChrysler AG

Commissariat à l'Énergie Atomique

Faurecia Industries

Volkswagen AG

Workpackage Leader:

Name:	Dr. Dariu Gavrilă
Company:	DaimlerChrysler AG
Phone:	+49 (731) 505 2388
Fax:	+49 (731) 505 4105
E-mail:	dariu.gavrila@daimlerchrysler.com

Project Coordinator:

Name:	Mr. Philippe MARCHAL
Company:	Faurecia Industries
Phone:	+33 (0)3 81 36 48 10
Fax:	+33 (0)3 81 36 45 20
E-mail:	pmarchal@seloncourt.faurecia.com

Contents

CONTENTS	4
1. EXECUTIVE SUMMARY	5
2. INTRODUCTION	6
3. DATA COLLECTION	7
3.1. Image Sensors	7
3.1.1. Standard Cameras (visible spectrum)	8
3.1.2. Infrared Cameras	8
3.2. Data Collection Methodology	9
3.2.1. Standard Cameras (visible spectrum)	10
3.2.2. Infrared Cameras	11
4. VRU IMAGE DATABASE	13
4.1. The Database at a Glance	13
4.2. Visible Spectrum Image Data Set	15
4.3. IR Image Data Set	15
4.3.1. VW Recordings	15
4.3.2. DC Recordings	21
5. ESTABLISHING “GROUND TRUTH”	23
6. CONCLUSIONS	27
7. REFERENCES	28

1. Executive Summary

This is the second part of the first Deliverable of the SAVE-U EU Project on vulnerable road user (VRU) protection, covering the collection and editing of sensor data. Image data was collected in both IR and visible light modality. The purpose of data collection was twofold: to obtain an indication of the sensor data quality allowing to select appropriate algorithms for pattern recognition, and furthermore, to compile a large image database. The latter would then facilitate statistical pattern recognition approaches to “learn” the image appearance of VRUs from examples, and allow for a systematic and quantitative evaluation of system performance, once “true” VRU image locations are labeled by a human operator (i.e. “ground truth”). We have compiled a huge image database containing over 14.000 images, making it - without doubt - the largest VRU database that has been collected worldwide. We conclude with a description of an interactive image-editing tool that was specially developed to support the VRU labeling effort. The SAVE-U image database will be made publicly available at the project web site www.save-u.org and is intended to promote research and collaboration at universities and research institutes EU- and world-wide on the important topic of VRU protection.

Keywords: Video, images, infra-red, data collection

2. Introduction

This is the second part of the first Deliverable of the **SENSORS AND SYSTEM ARCHITECTURE FOR VULNERABLE ROAD USERS PROTECTION (SAVE-U)** European Union (EU) project on sensor-based **Vulnerable Road User (VRU)** protection. It involves the collection and labeling of sensor data. Image data was collected in both IR and visible light modality. Section 3 describes the IR and visible light image sensors used for data collection and discusses the data collection methodology. The purpose of data collection was twofold, to obtain an impression of the sensor data quality allowing to select appropriate algorithms for pattern recognition, and furthermore, to compile a large image database. The latter would then facilitate statistical pattern recognition approaches to “learn” the image appearance of VRUs from examples, and allow for a systematic and quantitative evaluation of system performance, once “true” VRU image locations is labeled by a human operator (i.e. “ground truth”). We have compiled a huge image database containing over 14.000 images, making it - without doubt - the largest VRU database that has been collected worldwide. The database is described in Section 4. Section 5 presents an interactive image-editing tool that was developed to support the VRU labeling effort in order to obtain the important “ground truth”.

Finally, Sections 6 and 7 list the conclusions and references, respectively.

3. Data Collection

SAVE-U involves a sensor fusion platform combining radar and image sensors. Within the SAVE-U project plan it was acknowledged, however, that not all sensors (i.e. radar) would be available for the data collection as part of Work Package (WP) 1. It was nevertheless envisioned to at least start with the data collection for the individual image sensors as soon as possible, as part of WP 1, in order to guide the development of image processing algorithms.

Subsection 3.1 describes the image sensors used for data collection. Subsection 3.2 discusses the data collection methodology followed in WP 1 and provides an outlook to the important technical issues regarding the synchronization of various data sources for sensor fusion that will be addressed as part of the later Work Packages of SAVE-U.

3.1. Image Sensors

As part of WP 1, image data was collected with standard (visible spectrum) and IR cameras, see Subsections 3.1.1 and 3.1.2, respectively. The visible spectrum cameras were mounted inside the DaimlerChrysler vehicle behind the rear-view mirror, looking onto the road, see Figure 1. The IR cameras were either mounted on the roof (DaimlerChrysler) or attached to the front bumper (VW, see Figure 2) of the data collection test vehicle.



Figure 1: Video camera mounted on-board the DaimlerChrysler test vehicle.

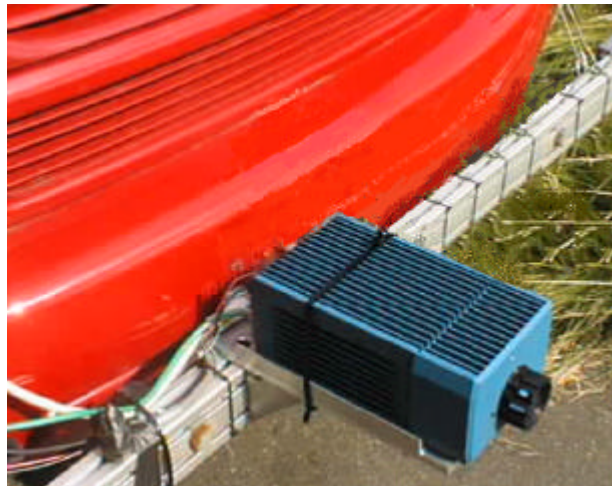


Figure 2: Infrared camera (CEDIP) attached to the front bumper of the VW test vehicle.

3.1.1. Standard Cameras (visible spectrum)

We used CCD cameras from manufacturers COBRA and PCO with image resolution 700x572 and 640x480 (“PAL resolution”), respectively, progressive scan. We considered both monocular and stereo configurations. The stereo configuration involved a base line of 30 cm with the cameras adjusted such that their optical axes were parallel and the horizontal scan lines were aligned.

For automotive application, high dynamic-range sensors will be needed to prevent blooming effects or dark spots due to local shadows. This can be achieved by sensors which show a logarithmic signal response with a dynamic range over 100 dB. Such a response can capture the large lighting variations seen at tunnels entrance/exits, at sunsets/sunrises or when facing headlights.

In the future, high-dynamic range CMOS cameras offer a promising alternative to CCD cameras due to their low-cost production for high volume and due to their further integration and miniaturization potential. In addition, on-chip image pre-processing will be feasible.

3.1.2. Infrared Cameras

Although IR video sensors are still too expensive for automotive application currently, their performance in bad weather conditions or at night makes them very attractive. That is the reason why SAVE-U has chosen to integrate and test an IR camera.

IR sensors based on new uncooled technology promise significant cost reductions and represent an important step forward towards the application of such sensors in the

automotive industry. Uncooled IR detectors enable cheaper cameras because they use focal plane arrays of microbolometers compatible with existing silicon fabrication lines.

Within the SAVE-U project, partner CEA has the task to define and simulate an uncooled IR detector optimized for pedestrian detection. One input for these simulations is images taken by existing IR sensors. The use of sensors based on CEA technology ensures a compatibility with the models of detector used in simulations. Therefore a commercially available camera “Jade UC” produced by CEDIP (www.cedip-infrared.com), a French company located near Paris, has been selected for the project. It has half-PAL resolution (320x240 pixels). The spectrum sensibility of the detector ranges from 8 μm to 12 μm which is well suited for human detection. This camera is equipped with a digital RS-422 output. The acquisition rate is 25 frames per second.



Figure 3: The CEDIP “Jade UC” uncooled IR camera

In addition to the CEDIP camera, a second uncooled IR camera was used for WP1 data collection, for comparison purposes. It is a prototype IR camera from AIM (AEG Infrared Module) with a microbolometer from BOEING, a half-PAL resolution (320x240 pixels) with a temperature sensitivity of 100m Kelvin.

3.2. Data Collection Methodology

Given the large disks of nowadays PCs, one no longer needs to record image data analog on VHS tape and digitize portions later off-line – with the associated loss of data quality. Instead, one can record images directly to disk by the appropriate acquisition boards placed in the PCI slots of the PC. Optical media (CD, DVD) can thereafter be used for archival purposes.

Acquisition boards must be chosen regarding the camera output formats (PAL, IEEE1394, etc.). For PC characteristics, hard drive interface (IDE or SCSI) has to be selected carefully considering the bandwidth required by video sources and the way the recording will be realized. Today, IDE hard drives can sustain a global bandwidth of 10 Mbytes/s for a single file access but this rate decreases dramatically for multiple file access. SCSI hard drives demonstrate access capacity on a single file over 20

Mbytes/s. In addition, SCSI is a smarter bus than IDE. It is capable of controlling the drives without any work by processors. A SCSI hard disk is therefore the preferred solution for data collection within the SAVE-U project, especially if PAL resolution is required.

Within the SAVE-U project, the data acquisition (processing) system must also eventually be able to deal with multiple video and radar sources, in parallel and distributed fashion. As these data sources have different acquisition (processing) rates, asynchronous communication is most suited for sensor fusion. It is envisioned that each acquisition (processing) unit will submit its data to a "fusion" engine as soon as it becomes available, including local time stamp. The system clocks of the various units are aligned by periodical synchronization.

In terms of communication channels and protocols, since the vehicle CAN-bus provides insufficient throughput for image data transfer at half-PAL resolution, Fast Ethernet is the only other viable alternative.

In terms of protocols, an attractive solution is the Parallel Virtual Machine (**PVM**, http://www.epm.ornl.gov/pvm/pvm_home.html), that permits a heterogeneous collection of Unix and/or NT computers hooked together by a network to be used as a single large parallel computer. The software, presenting a convenient interface on top of TCP/IP, is very portable and is available free. Hundreds of sites around the world are using PVM to solve scientific, industrial, and medical problems, in addition to PVM's use as an educational tool to teach parallel programming. With tens of thousands of users, PVM has become a de-facto standard for distributed computing worldwide.

In this context, one can also mention the commercial software package **RTMaps** from the French company Intempora (www.intempora.com). Originally developed for the automotive industry, it can record multiple sensors data (video, CAN bus...) in real-time in order to constitute a database of scenarios. Every data is dated accurately when they enter the system and this information is inserted into the database. Then, scenarios can be re-played with perfect data synchronization. RTMaps has furthermore facilities to visualize and analyze data, and it is distributed with some basic pattern recognition algorithms already implemented, which can be extended by modules developed by the user. A developer license for RTMaps costs about €20.000.

3.2.1. Standard Cameras (visible spectrum)

We used the ITI IMPCI and PCO frame-grabbers to store image data digitally onto disk. Our PC system had an integrated SCSI controller. Existing DaimlerChrysler application software accesses the underlying frame-grabbing routines and offers different modes of

saving data. In the “continuous” mode, image data is saved to disk without pause, at a rate limited by the SCSI disk subsystem. In the “buffer” mode, upon a user-provided trigger, the contents of a ring-buffer residing in PC internal memory are flushed. This allows obtaining recorded data at the full acquisition rate, although only during a limited time (i.e. depending on size of buffer).

3.2.2. Infrared Cameras

The “Jade UC” camera system is composed of a power supply module, an RS 232 cable, a numerical video cable and an acquisition board plugged into a PC slot.

The power supply module generates 24 VDC from any AC source ranging from 100 VAC to 250 VAC. To fix camera parameters, a RS232 communication cable is used to connect the camera to a PC. In that case, an application running on a PC (Cirrus software) enables to change these parameters, see Figure 4 The acquisition board performs image grabbing with 14-bits of dynamic under the control of the Recorder application.

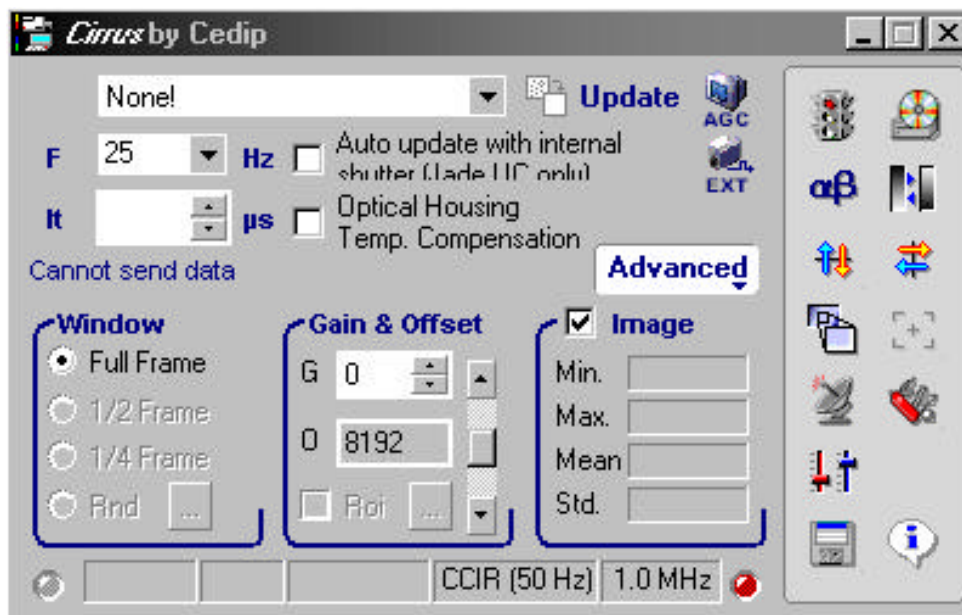


Figure 4: Cirrus Application for parameter settings

One of the main parameters is the Non Uniform Correction (NUC) table. The sensing elements of the camera are affected by spatial and temporal noises. This table brings a correction to cope with their non uniform response. For that goal, a calibration phase prior to the recording is necessary. This phase is done by placing one or two points at different temperatures in front of the camera lens. Then, gain and offset coefficients are calculated and averaged on some images.

The “gain & offset” field of the Cirrus interface is only a display facility. The coefficients are used to transform 14-bits images into 8-bits images but only for display purpose. These parameters do not affect the value of the 14-bits data taken by the digital grabbing but only interact with the analog video output.

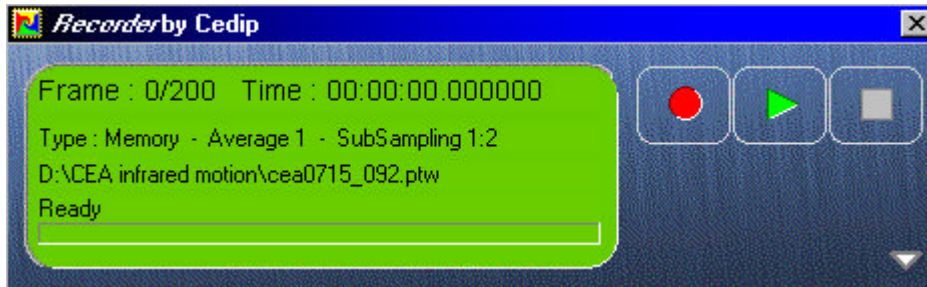


Figure 5: Recorder application for numerical recording

The Recorder application (see Figure 5) runs the acquisition board and enables to start and stop a recording either by mouse clicking or by setting a grabbing period in seconds. The video output file is in a CEDIP proprietary format (.ptw file) but CEA has developed a software to convert this format in a sequence of 16 bit-TIFF images.

4. VRU Image Database

This section describes the SAVE-U VRU image database, compiled following the sensors and data collection methodology introduced in previous section. The VRU database can be considered unique worldwide, because of its huge size: it contains more than 14.000 images and 180 sequences.

The benefits of a large database are twofold. First, it provides a wealth of training data for statistical pattern matching techniques. These “learn” the VRU appearance from examples; which is important, since good prior, explicit models are hard to define. Second, it allows to evaluate system performance on a truly large dataset, so that the results can be considered representative of the true physical traffic situation.

The SAVE-U image database will be made publicly available at the project web site www.save-u.org and is intended to promote research and collaboration at universities and research institutes EU- and world-wide on the important topic of VRU protection.

4.1. The Database at a Glance

The database contains about 14.000 images and 180 sequences, derived from the standard visible-light cameras (Subsection 3.1.1) or from the infrared sensors (Subsection 3.2.2).

Table 1 provides a break-down of image type versus the source providing the data. Given the wide pre-existing availability of standard video data, most of the data collection effort was directed towards the IR modality.

	IR Spectrum Images (Sequences)	Visible Spectrum Images (Sequences)
VW	6250 (52)	
DC	6139 (109)	1972 (20)

Table 1: The SAVE-U Database Overview

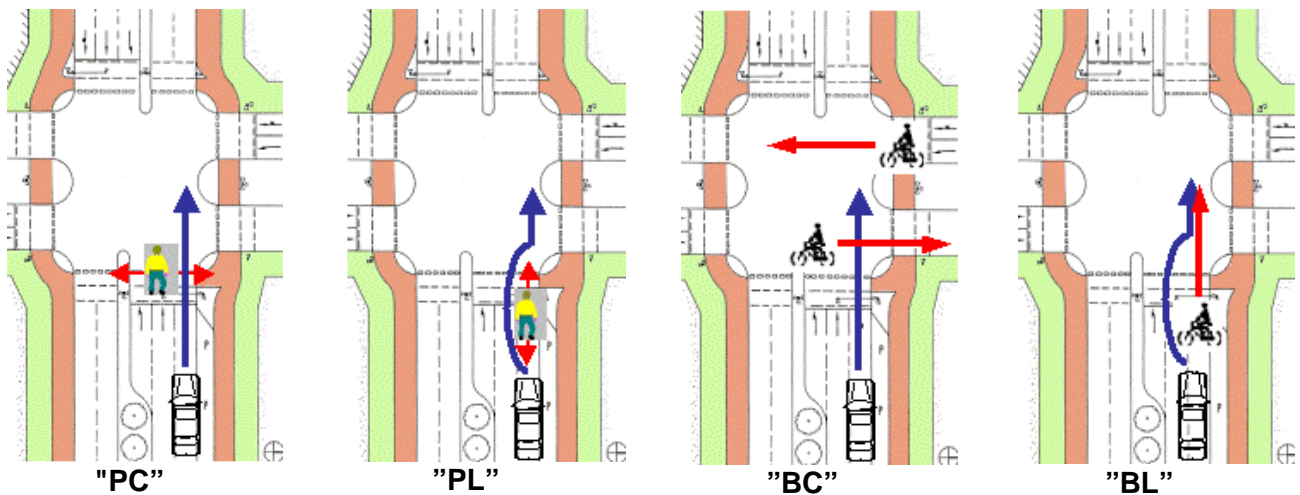


Figure 6: Four different SAVE-U scenarios

Terminology "XY":

„X“ is either „P“ (Pedestrian) or „B“ (Bicyclist)

„Y“ is either „C“ (Crossing) or „L“ (Longitudinal movement)

The image database is structured according to the scenarios shown in Figure 6¹. This classification is solely based on the relative geometry between vehicle and the VRU. Thus, for example, although Figure 6 shows an intersection, scenario "PC" for example covers both the case of a pedestrian walking on a marked pedestrian crossing ("zebra") at an intersection and the case of a pedestrian crossing the road at an arbitrary, unmarked location. Pedestrians can be both adults and children and can be walking at various speeds (i.e. from running to standing still).

Note that the SAVE-U scenarios shown in Figure 6 are by no means exhaustive, they only involve a straight or slightly curved road where a VRU appears in front of the vehicle. To account for all VRU-vehicle cases not covered by Figure 6 (e.g. turn at an intersection), an additional class was introduced for database organization. One further class involved image data which did not contain VRUs at all.

The images in the SAVE-U VRU image database are stored either in the „tagged image file format“ (TIFF) or in the “portable greymap file format” (PGM). The TIFF format has the advantage that it can handle images with more than 8 bit/pixel, whereas the PGM format is a bit faster to access, because of its uncompressed nature. Most of the images have a “half-PAL” resolution, of 320 x 240 pixels, which, from previous experience, appears sufficient to perform image-based VRU detection and classification.

A subset of the data is annotated with „ground truth“, the true locations of VRUs in the images. Presently, this was done by drawing in the object contours manually into

¹ SAVE-U partner Volkswagen in fact categorized their infrared image data set by the accident types of the German Insurance Association (GDV), but a straightforward mapping exists onto the SAVE-U scenarios.

images. A tool for semiautomatic image labeling has recently been completed and it is described in Section 5.

The SAVE-U VRU database is distributed by means of two DVD-R disks, which contain the data collected by SAVE-U partners Volkswagen and DaimlerChrysler.

4.2. Visible Spectrum Image Data Set

The image data set recorded with the standard visible light cameras contains 15 mono and 5 stereo sequences, see Table 2. Most of them contain pedestrians in urban environment walking longitudinal or across the street and are instances of scenarios PC and PL. The remaining three sequences contain more complex scenarios do not fit directly to any of the defined scenarios.

SAVE-U Scenario	MONO IMAGES Images (Sequences)	STEREO IMAGES Images (Sequences)	Sum Images (Sequences)
PC	210 (6)	954 (4)	1164 (10)
PL	226 (6)	350 (1)	576 (7)
OTHERS	232 (3)		232 (3)
Sum	668 (15)	1304 (5)	1972 vi

Table 2: Overview of Visible Spectrum Image Data Set

4.3. IR Image Data Set

4.3.1. VW Recordings

The following table contains details about the infrared image subset that was recorded by SAVE-U partner VOLKSWAGEN (VW). In total 6250 images (3,9 GByte of data) were recorded in 12 different types of traffic situations. The selection of the recorded traffic situations was influenced by the top ten accident types in Germany, as discussed in SAVE-U Deliverable 1-A [GAVRILA03] (for the accident classification by the German Insurance Association, GDV, see also <http://www.gdv-isk.de/downloads/unfalltypen.zip>). In addition, a number of country road and other traffic situations were recorded.

To have a wide spectrum of environmental conditions in the infrared images in each type of situations several instances were observed and stored. The images include pedestrians, bicyclists as well as motorcycles. The recordings were performed in the late afternoon on a typical summer day.

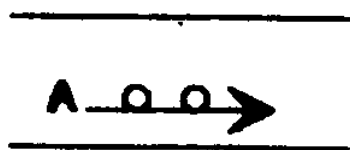
Type of Scenario (similar to Type of Accident)	Number of recorded Situations	Total number of recorded infrared images
101	2	350
141	9	1650
221	1	100
401	11	1100
411	2	200
421	14	1400
423	5	650
424	3	300
431	2	200
441	1	100
451	1	100
461	1	100
Total		6250

Table 3: Overview of recorded scenario type – VW

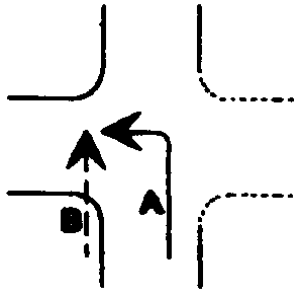
4.3.1.1. Configuration No. 101: Curve left



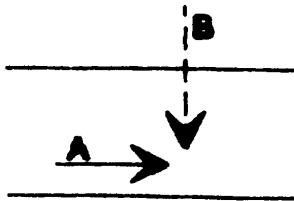
4.3.1.2. Configuration No. 141: Straight Ahead



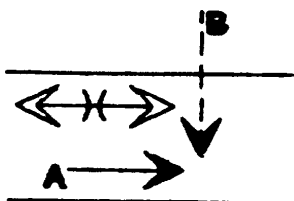
4.3.1.3. Configuration No. 221: Turn to the left, Pedestrian coming from separate sidewalk



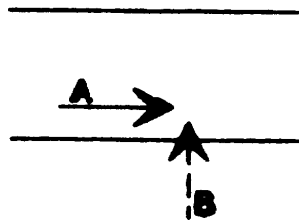
4.3.1.4. Configuration No. 401: VRU coming from the left



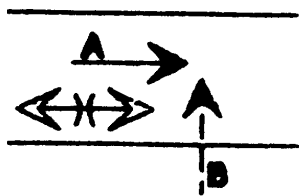
4.3.1.5. Configuration No. 411: Pedestrian coming from left, reduced visibility



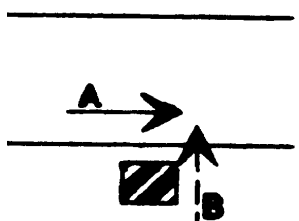
4.3.1.6. Configuration No. 421 Pedestrian coming from right



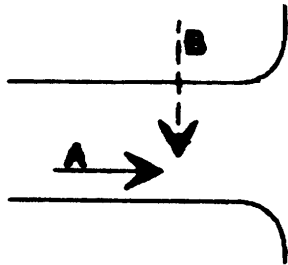
4.3.1.7. Configuration No. 423: Pedestrian coming from right, reduced visibility



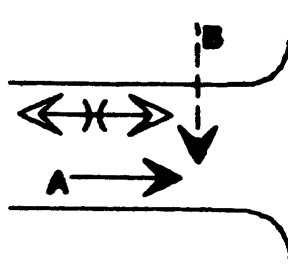
4.3.1.8. Configuration No. 424: Pedestrian coming from right, reduced visibility



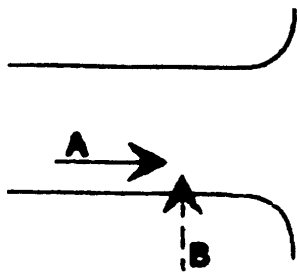
4.3.1.9. Configuration No. 431: Pedestrian coming from left



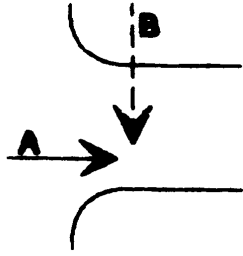
4.3.1.10. Configuration No. 441: Pedestrian coming from the left, reduced visibility



4.3.1.11. Configuration No 451: Pedestrian coming from the right



4.3.1.12. Configuration No. 461: Curve left



4.3.1.13. Configuration "Not Specified"



No Pictogram

4.3.1.14. Configuration "No Pedestrians"



No Pictogram

4.3.2. DC Recordings

Table 4 gives an overview of the infrared data set recorded by DaimlerChrysler. In sum, it contains more than 6000 images belonging to more than 100 sequences. Sample images taken with the AIM camera of the different SAVE-U scenarios PC, PL, BC and BL, defined in Section 4.1, are shown in Figure 7 and Figure 8.

SAVE-U Scenario	CEDIP-CAMERA DAY Images (Sequences)	CEDIP-CAMERA NIGHT Images (Sequences)	AIM-CAMERA ² DAY Images (Sequences)	Sum Images (Sequences)
PC	1500 (15)		279 (13)	1779 (28)
PL	100 (1)	300 (3)	335 (13)	735 (17)
BC	300 (3)		4 (1)	304 (4)
BL	300 (3)		52 (3)	352 (6)
NO VRU	400 (4)	600 (6)	24 (1)	1024 (11)
OTHERS	600 (6)	1000 (10)	345 (27)	1945 (43)
Sum	3200 (32)	1900 (19)	1039 (58)	6139 (109)

Table 4: Overview of recorded scenarios with the CEDIP camera



**Figure 7: Sample images from scenarios PC and PL
(Pedestrian, Crossing/Longitudinal movement)**

² For this dataset “ground truth” data is available, see Section 5.



Figure 8: Sample images from scenarios BC and BL (Bicyclist, Crossing/Longitudinal movement)

5. Establishing “Ground Truth”

In order to train statistical pattern matching methods, or to measure system performance during testing, we need to know the “true” position and spatial extent of the VRU in the images of the database, i.e. the “ground truth”. To establish this, a semiautomatic labeling-tool called **VisiCurve** has been developed by the SAVE-U partner DaimlerChrysler, which assists the user in outlining the VRU object contours in images and in establishing temporal correspondence across the images of a sequence. See Figure 9.

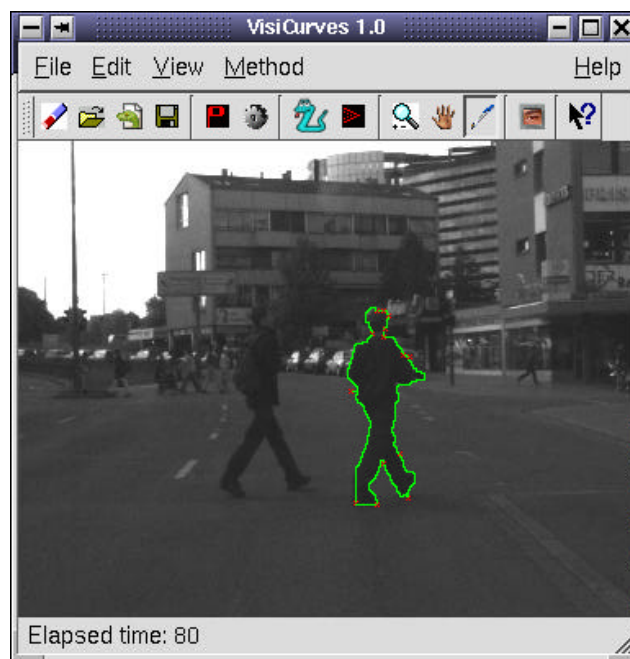


Figure 9: The VisiCurve Tool for Computer-Assisted Image Labeling – Main Window

The main aim of VisiCurve is to avoid the need for a user having to specify pixel-by-pixel the entire object contour. This “manual” contour labeling is not only very time consuming but also a very tedious. Computer vision techniques for segmentation can in principle support the labeling process, since they are designed to find object boundaries automatically. They usually utilize a certain degree of prior knowledge on object appearance and lock on low level image features. The algorithms typically involve an object model, whose parameters can be restricted by the prior knowledge to certain ranges of valid object configurations. Various optimization and estimation schemes can be applied to adjust the parameters of the model in a way to yield a good fit between the resulting object and image features. Most of them require a user supplied initial guess of the parameter set.

VisiCurve supports three different “power-assist” techniques³ :

1. **The classical Snake algorithm** [KASS88]: A snake is a spline curve, whose propagation depends on image forces, which pull it towards image features, and internal forces, which enforce curve smoothness and continuity. Once the snake is initialized by the user sufficiently close, it “crawls” iteratively towards the VRU boundary.
2. **Hermite Deformable Contours** [GAVRILA96]: Hermite deformable contours are also energy minimizing cubic spline curves, considering internal and external forces. They are represented by a set of parameters: control points with associated tangent vectors. After they have been initialized by the user, these parameters are optimized by dynamic programming, a sequential pattern matching technique, which finds the optimal solution with regard to the defined energy formulation.
3. **Minimal Path Approach** [COHEN93]: The minimal path approach finds the global minimum of a contour’s energy between two user supplied points. As above, the energy function contains an internal regularization term and an external potential. A useful interactive variant of this method consists of the system continuously drawing the best path between a previously confirmed “control” point and the current location of the mouse on top of the image.

This in addition to the

4. **Freehand (“Manual”) Approach.** Here the exact contour pixels are recorded that lie under the mouse cursor which is moved by the user along the object contour.

VisiCurve offers various zoom in/out and translate capabilities to allow a better focus on the relevant portions of an image. Once an object is labeled by one of the above three power-assist methods or manually, it is further handled by the object window, see Figure 10. If the labeled object was not labeled in a previous image of a sequence, the user designates it as “New” and specifies its object type (e.g. “pedestrian”, “vehicles”). An image thumbnail of the new object appears in the corresponding category in the Object Window and the system internally generates a new object id. If however, the labeled object already appeared in a previous image of a sequence, it is desirable to maintain temporal correspondence. This is conveniently done by the user clicking on the corresponding image thumbnail in the VisiCurve Object Window. The system internally assigns then the correct object id to the labeled object. The use of the object

³ Implementations for methods 1-3 were adapted from software from INTEL’s Open CV Library , Gavrilas’s Ph.D. Thesis at the University of Maryland and a DaimlerChrysler project on industrial vision, respectively.

thumbnails in the Object Window has the advantage that the bookkeeping of the object ids remains transparent for the user. When all objects in an image have been annotated, the “ok” button serves to save the image labels to a corresponding data file.

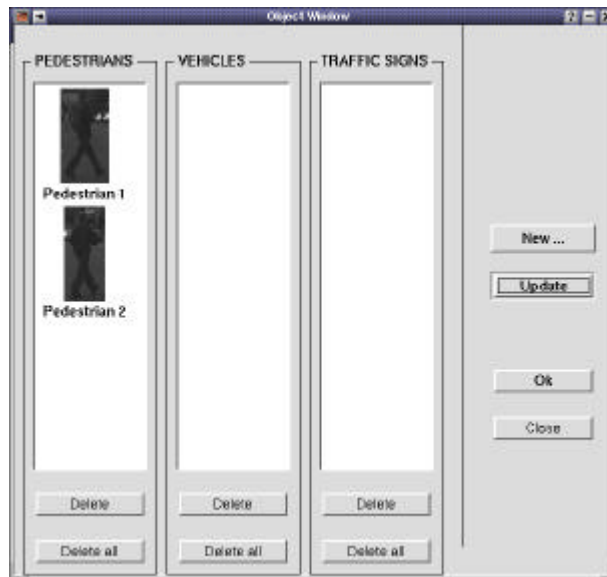


Figure 10: VisiCurve’s Object Window

The resulting data file (Figure 11) contains the classes, bounding boxes, and contour points for each labeled object.

```

Object Id: 4
Object Label: ped
Object Category: Object_FV
Object Bounding Box: 189 367 240 382
Object No Points: 148
Object Contour: 379 190 379 191 379 192
379 193 379 194 379 195 379 196 380 196
369 204 369 203 370 203 370 202 370 201
374 197 374 196 373 196 372 196 372 195
382 214 382 215 381 215 381 216 381 217
379 223 379 224 379 225 379 226 379 227
369 205 368 205 382 212 382 211 382 210
382 202 368 205 369 225 369 226 369 227
368 208 368 209 368 210 368 211 368 212
369 219 369 220 369 221 369 222 369 223
369 230 369 231 369 232 369 233 369 234
367 240 368 240 369 240 370 240 371 240
379 236 377 240 378 240 378 239 378 238
    
```

Figure 11: Resulting Data File with Ground Truth Data

With this information, one can for example extract for each VRU object in the database the rectangular $W \times H$ subimage corresponding to its bounding box. The subimage can then be mapped onto a feature vector by several techniques, most easily by aggregating the pixel intensities into a $N = W \times H$ dimensional feature vector. Typically, one thereafter applies a dimensionality reduction technique such as Principal Component Analysis (PCA) leading to a lower dimensional $N' \ll N$ feature vector. The resulting feature vectors can be used in standard pattern classification approaches

(Neural Networks, Polynome Classifiers, Support Vector Machines, etc) in order to “learn” the appearance of a particular VRU object class.

Alternatively, the availability of the bounding box information of “true” image location of the VRU in a test image database can be used to quantify system detection performance. Detected solutions can, for example, be counted as “correct” if the corresponding bounding boxes lie within a user-specified tolerance on position, box width and height, compared to the bounding boxes of the labeled solutions. Several small programs have been developed for this purposes.

Compared to bounding boxes, the availability of precise contour data allows a more detailed description of a VRU object in the image, and can thus support more powerful feature vector representations for object learning. If actual VRU pose (e.g. pedestrian feet closed vs. feet apart) is a test criterion as well, labeled contour data is also essential for the system evaluation phase.

Above-mentioned techniques for image-based VRU object learning and system evaluation will be pursued in SAVE-U as part of Work Package 6. At that point, we expect VisiCurve to play an important role. In order to get a head start, simultaneous with the development of VisiCurve, SAVE-U partner DaimlerChrysler, had the VRU object contours in AIM-camera IR dataset labeled “manually” (1039 images, 58 sequences). See corresponding DVD-R of the SAVE-U VRU database.

We conclude this section by mentioning that VisiCurve is implemented in C++ and currently runs under the Linux operating system. The graphical user interface is based on the cross-platform C++ Qt library (www.trolltech.no), which facilitates the porting VisiCurve to a Windows OS, if desired.

6. Conclusions

The second part of the first Deliverable of the SAVE-U EU Project on vulnerable road user (VRU) protection, covering data collection and editing, was successfully concluded.

A huge and unique VRU database, containing over 14.000 images in visible and IR spectrum, was collected and organized according to the SAVE-U scenarios. A special interactive image-editing tool was developed to support the object labeling tool; part of the database was already processed. The SAVE-U image database will be made publicly available at the project web site www.save-u.org and is intended to promote research and collaboration at universities and research institutes EU- and world-wide on the important topic of VRU protection.

7. References

- **[COHEN93]** L. Cohen and R. Kimmel, “Global Minimum for Active Contour Models: A Minimal Path Approach”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.15, nr.11, pp.1131-1147, 1993.
- **[GAVRILA96]** D. M. Gavrilă, “Hermite Deformable Contours”, *Proc. of IEEE International Conference on Pattern Recognition*, pp. 130-135, Vienna, Austria, 1996.
- **[GAVRILA03]** D. M. Gavrilă, P. Marchal and M.-M. Meinecke, “Vulnerable Road User Scenario Analysis”, SAVE-U Deliverable 1A, 2003.
- **[KASS88]** M. Kass, A. Witkin and D. Terzopoulos, “Snakes: Active Contour Models”, *Int. J. Computer Vision*, vol 1, pp. 321-331, 1988.